Figure 1

Schematic layout of the arrangement of the genetic locus encoding the signal peptide precursor, the histidine kinase and the response regulator. Note that this arrangement is different from other loci in related streptococci for the following reasons: a) The comC gene is transcribed from its own promoter alone, unlike the genes thus far described in other streptococci that are arranged in an operon-like cluster with the comC/DE genes being transcribed from a single promoter.

b) The comC gene is separated from the *comD* gene by 148 nucleotides.

Streptococcus mutans
***ComCDE* Operon**

Figure 2
Sequences of the open reading frames encoding the signal peptide precursor (ComC), the histidine kinase (ComD), and the response regulator (ComE).

➢ S. mutans comC gene
Encodes a precursor to a signal peptide
[ATGAAAAAAACACTATCATTAAAAAATGACTTTAAAGAAATTAAGACTGATGAATTAG
AGATTATCATTGGCGGA(AGCGGAAGCCTATCAACATTTTTCCGGCTGTTTAACAGAAG
TTTTACACAAGCTTTGGGAAAA)]TAA

➢ S. mutans CSP encoding sequence
Competence Signal Peptide
AGCGGAAGCCTATCAACATTTTTCCGGCTGTTTAACAGAAGTTTTACACAAGCTTTGGG
AAAA [SEQ ID NO:1]

➢ S. mutans comD gene
Encodes a protein that functions as a histidine kinase receptor
[ATGAATGAAGCCTTAATGATACTTTCAAATGGTTTATTAACTTATCTAACCGTTCTAT
TTCTCTTGTTTCTATTTTCTAAGGTAAGTAATGTCACTTTATCGAAAAAGGAATTAACT
CTTTTTTCGATAAGCAATTTTCTGATAATGATTGCTGTTACGATGGTGAACGTAAACCT
GTTTTATCCTGCAGAGCCTCTTTATTTTATAGCTTTATCAATTTATCTTAATAGACAGA
ATAGTCTTTCTCTAAATATATTTTATGGTCTGCTGCCTGTTGCCAGTTCTGACTTGTTT
AGGCGGGCAATCATATTCTTTATCTTGGATGGAACTCAAGGAATTGTAATGGGCAGTAG
CATTATAACCACCTATATGATCGAGTTTGCAGGAATAGCGCTAAGTTACCTCTTTCTCA
GTGTGTTCAATGTTGATATTGGTCGACTTAAAGATAGTTTGACCAAGATGAAGGTCAAA
AAACGCTTGATTCCAATGAATATTACTATGCTTCTATACTACCTTTTAATACAGGTATT
GTATGTTATAGAGAGTTATAATGTGATACCGACTTTAAAATTTCGTAAATTTGTCGTTA
TTGTCTATCTTATTTTATTTTTGATTCTGATCTCATTTTTAAGCCAATATACCAAACAA
AAGGTTCAAAATGAGATAATGGCACAAAAGGAAGCTCAGATTCGAAATATCACCCAGTA
TAGTCAGCAAATAGAATCTCTTTACAAGGATATTCGAAGTTTCCGCCATGATTATCTGA
ATATTTTAACTAGCCTCAGATTAGGCATTGAAAATAAAGATTTAGCTAGTATTGAAAAG
ATTTACCATCAAATCTTAGAAAAAACAGGACATCAATTGCAGGATACCCGTTATAATAT
CGGCCATCTAGCTAATATTCAAAACGATGCTGTCAAGGGTATCTTGTCAGCAAAAATCT
TAGAAGCTCAGAATAAAAAGATTGCTGTCAATGTAGAAGTCTCAAGTAAAATACAACTG
CCTGAGATGGAGTTGCTTGATTTCATTACCATACTTTCTATCTTGTGTGATAATGCCAT
TGAGGCTGCTTTCGAATCATTAAATCCTGAAATTCAGTTAGCCTTTTTTAAGAAAAATG
GCAGTATAGTCTTTATCATTCAGAATTCCACCAAAGAAAAACAAATAGATGTGAGTAAA
ATTTTTAAAGAAAACTATTCCACTAAAGGCTCCAATCGCGGTATTGGTTTAGCAAAGGT
GAATCATATTCTTGAACATTATCCCAAAACCAGTTTACAAACAAGCAATCATCATCATT
TATTCAAGCAACTCCTAATAATAAAA]TAG

➢ S. mutans comE gene
Encodes a response regulator that activates transcription of a number of genes

[ATGATTTCTATTTTTGTATTGGAAGATGATTTTTTACAACAAGGACGTCTTGAAACCA
CCATTGCAGCTATCATGAAAGAAAAAAATTGGTCTTATAAAGAATTGACTATTTTTGGA
AAACCACAACAACTTATTGACGCTATCCCTGAAAAGGGCAATCACCAGATTTTCTTTTT
GGATATTGAAATCAAAAAAGAGGAAAAGAAAGGACTGGAAGTAGCCAATCAGATTAGAC
AGCATAATCCTAGTGCAGTTATTGTCTTTGTCACGACACATTCTGAGTTTATGCCCCTC
ACTTTTCAGTATCAGGTATCTGCTTTGGATTTTATTGATAAATCTTTGAATCCTGAGGA
GTTCTCCCACCGCATTGAATCAGCGCTGTATTATGCTATGGAAAACAGCCAGAAGAATG
GTCAATCAGAGGAACTTTTTATTTTCCATTCATCTGAAACTCAGTTTCAGGTCCCTTTT
GCTGAGATTCTGTATTTTGAAACATCTTCAACAGCCCATAAGCTCTGCCTTTATACTTA
TGATGAACGGATTGAATTCTACGGCAGTATGACTGACATTGTTAAAATGGATAAGAGAC
TTTTTCAGTGCCATCGCTCTTTTATTGTCAATCCTGCCAATATTACCCGTATTGATCGG
AAAAAACGCTTGGCCTATTTTCGAAATAATAAGTCTTGTCTTATTTCACGAACTAAGTT
AACAAAACTGAGAGCTGTGATTGCTGATCAAAGGAGAGCAAAA] TGA

Figure 3
The amino acid sequences of the signal peptide precursor (ComC), the histidine kinase (ComD), and the response regulator (ComE).

> S. mutans ComC protein (CSP Precursor)
MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK

> S. mutans ComD protein (Histidine Kinase)
MNEALMILSNGLLTYLTVLFLLFLFSKVSNVTLSKKELTLFSISNFLIMIAVTMVNVNL
FYPAEPLYFIALSIYLNRQNSLSLNIFYGLLPVASSDLFRRAIIFFILDGTQGIVMGSS
IITTYMIEFAGIALSYLFLSVFNVDIGRLKDSLTKMKVKKRLIPMNITMLLYYLLIQVL
YVIESYNVIPTLKFRKFVVIVYLILFLILISFLSQYTKQKVQNEIMAQKEAQIRNITQY
SQQIESLYKDIRSFRHDYLNILTSLRLGIENKDLASIEKIYHQILEKTGHQLQDTRYNI
GHLANIQNDAVKGILSAKILEAQNKKIAVNVEVSSKIQLPEMELLDFITILSILCDNAI
EAAFESLNPEIQLAFFKKNGSIVFIIQNSTKEKQIDVSKIFKENYSTKGSNRGIGLAKV
NHILEHYPKTSLQTSNHHHLFKQLLIIK

> S. mutans ComE protein (Response Regulator)
MISIFVLEDDFLQQGRLETTIAAIMKEKNWSYKELTIFGKPQQLIDAIPEKGNHQIFFL
DIEIKKEEKKGLEVANQIRQHNPSAVIVFVTTHSEFMPLTFQYQVSALDFIDKSLNPEE
FSHRIESALYYAMENSQKNGQSEELFIFHSSETQFQVPFAEILYFETSSTAHKLCLYTY
DERIEFYGSMTDIVKMDKRLFQCHRSFIVNPANITRIDRKKRLAYFRNNKSCLISRTKL
TKLRAVIADQRRAK

Figure 4
The deduced amino acid sequence of the signal peptide precursor in various strains and its predicted cleavage site. The original peptide is expressed as a 46-amino acid peptide that is cleaved after the glycine-glycine residues to generate an active signal peptide.

```
BM71 CSP      1 MKKTPSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
GB14 CSP      1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
H7 CSP        1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
JH1005 CSP    1 MKKTLSLKNDFKEIKTDELEIIIGGSGTLSTFFRLFNRSFTQA    43
LT11 CSP      1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
NG8 CSP       1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
UAB159 CSP    1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
                **** ********************.***************

consensus: 1 MKKTLSLKNDFKEIKTDELEIIIGG SGSLSTFFRLFNRSFTQALGK 46
predicted cleavage site:              ^
```

Figure 5
The synthetic signal peptide that is effective at inducing competence, biofilm formation and acid tolerance in *Streptococcus mutans*.

```
SGSLSTFFRLFNRSFTQALGK [SEQ ID NO:2]
```

Figure 6
The natural activity of the signal/receptor system functioning *in vitro* in model biofilms as determined by the ability of various strains of *S. mutans* to accept donor plasmid DNA conferring erythromycin resistance.

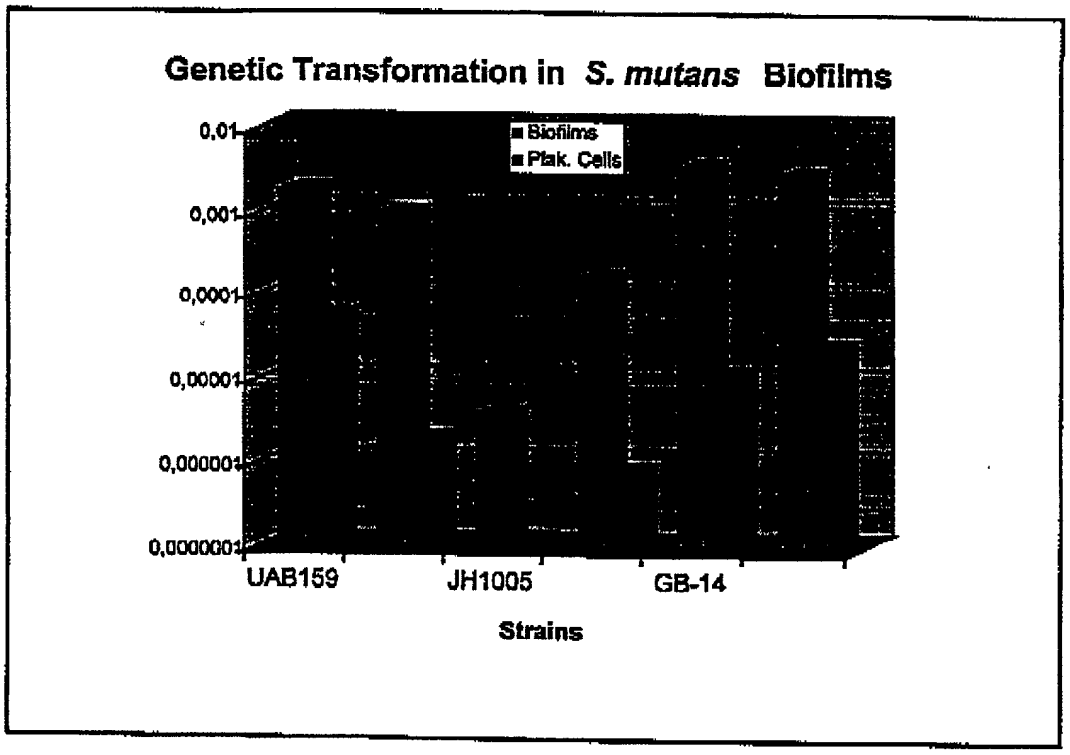


Figure 7
Induction of genetic transformation in *Streptococcus mutans* by synthetic competence stimulating peptide (SCSP)[1]

Strain	Peptide added Number of Transformants/Recipients	No peptide Number of Transformants/Recipients
UAB15	4.65×10^{-1}	1.78×10^{-6}
JH1005[2]	6.98×10^{-2}	0

[1]The final concentration of SCSP used was 500 ng/ml.
The strain contains a nonsense mutation in the *comC* gene encoding the CSP.

Figure 8
List of the primers used to amplify the genes or internal regions of the target genes by polymerase chain reaction (PCR) for subsequent sequencing or inactivation.

ComC region

ComC Primer Pair: F5-B5

[F5] 23406-23424 5'- AGTTTTTTGTCTGGCTGCG -3'
19 nt forward primer
pct G+C: 47.4 Tm: 50.5

[B5] 24056-24037 5'- TCCACTAAAGGCTCCAATCG -3'
20 nt backward primer
pct G+C: 50.0 Tm: 51.9

651 nt product for F5-B5 pair (23406-24056)
Optimal annealing temp: 50.3
pct G+C: 30.9 Tm: 71.5

ComD region

ComD Primer Pair: F1-B1

[F1] 392-415 5'- CGCTAAGTTACCTCTTTCTCAGTG -3'
24 nt forward primer
pct G+C: 45.8 Tm: 51.6

[B1] 683-663 5'- GCTTCCTTTTGTGCCATTATC -3'
21 nt backward primer
pct G+C: 42.9 Tm: 50.8

292 nt product for F1-B1 pair (392-683)
Optimal annealing temp: 49.5
pct G+C: 30.8 Tm: 70.2

ComE region

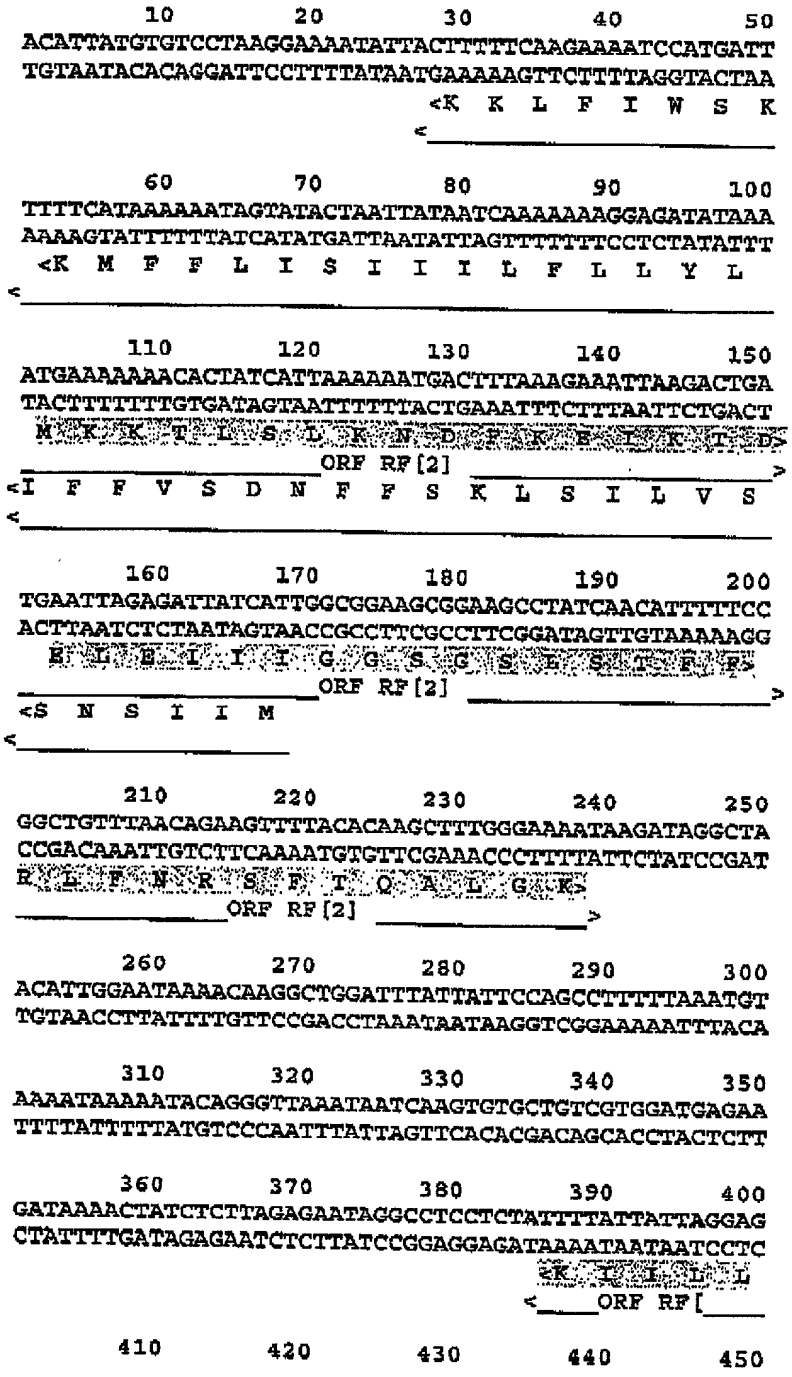
ComE Primer Pair: F1-B1

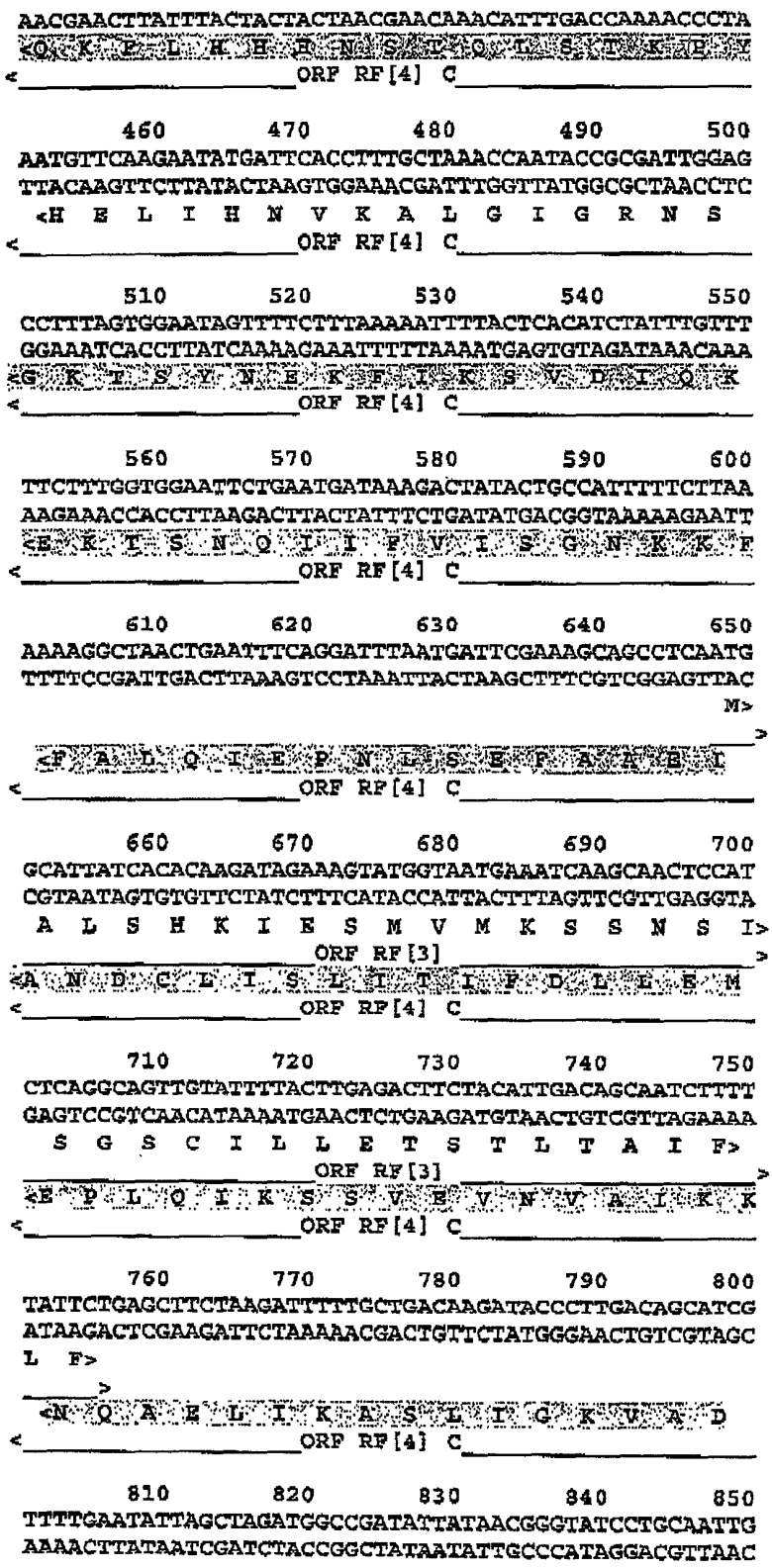
[F1] 145-165 5'- CCTGAAAAGGGCAATCACCAG -3'
21 nt forward primer
pct G+C: 52.4 Tm: 55.9

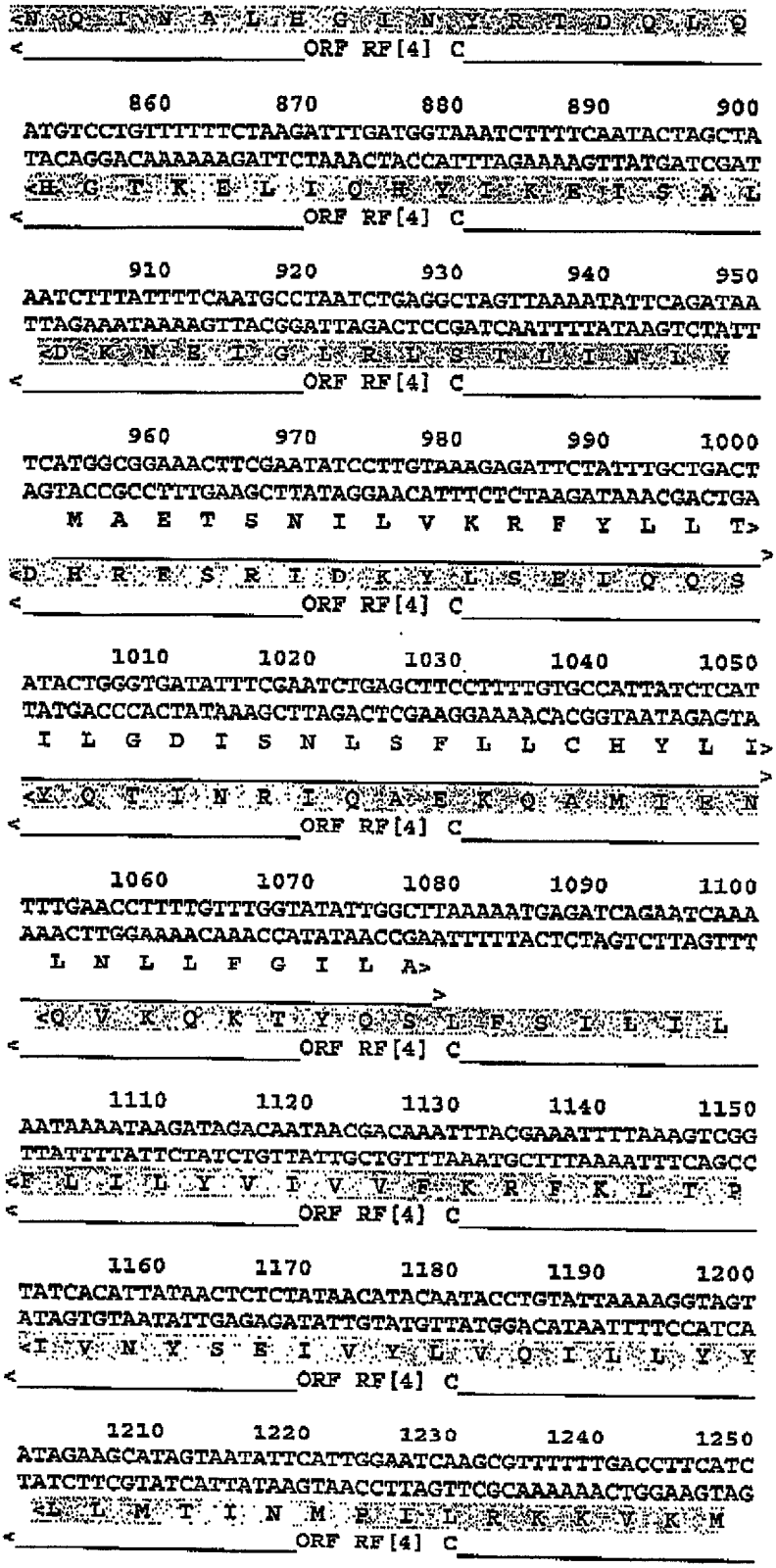
[B1] 606-585 5'- GCGATGGCACTGAAAAAGTCTC -3'
22 nt backward primer
pct G+C: 50.0 Tm: 55.4

462 nt product for F1-B1 pair (145-606)
Optimal annealing temp: 53.6
pct G+C: 38.3 Tm: 74.1

Figure 9
ComCDE local region. The ComC (first highlighted region; nucleotides 101 to 241), ComD (second highlighted region; nucleotides 383 to 1708) and ComE (third highlighted region; nucleotides 1705 to 2457) proteins are highlighted.

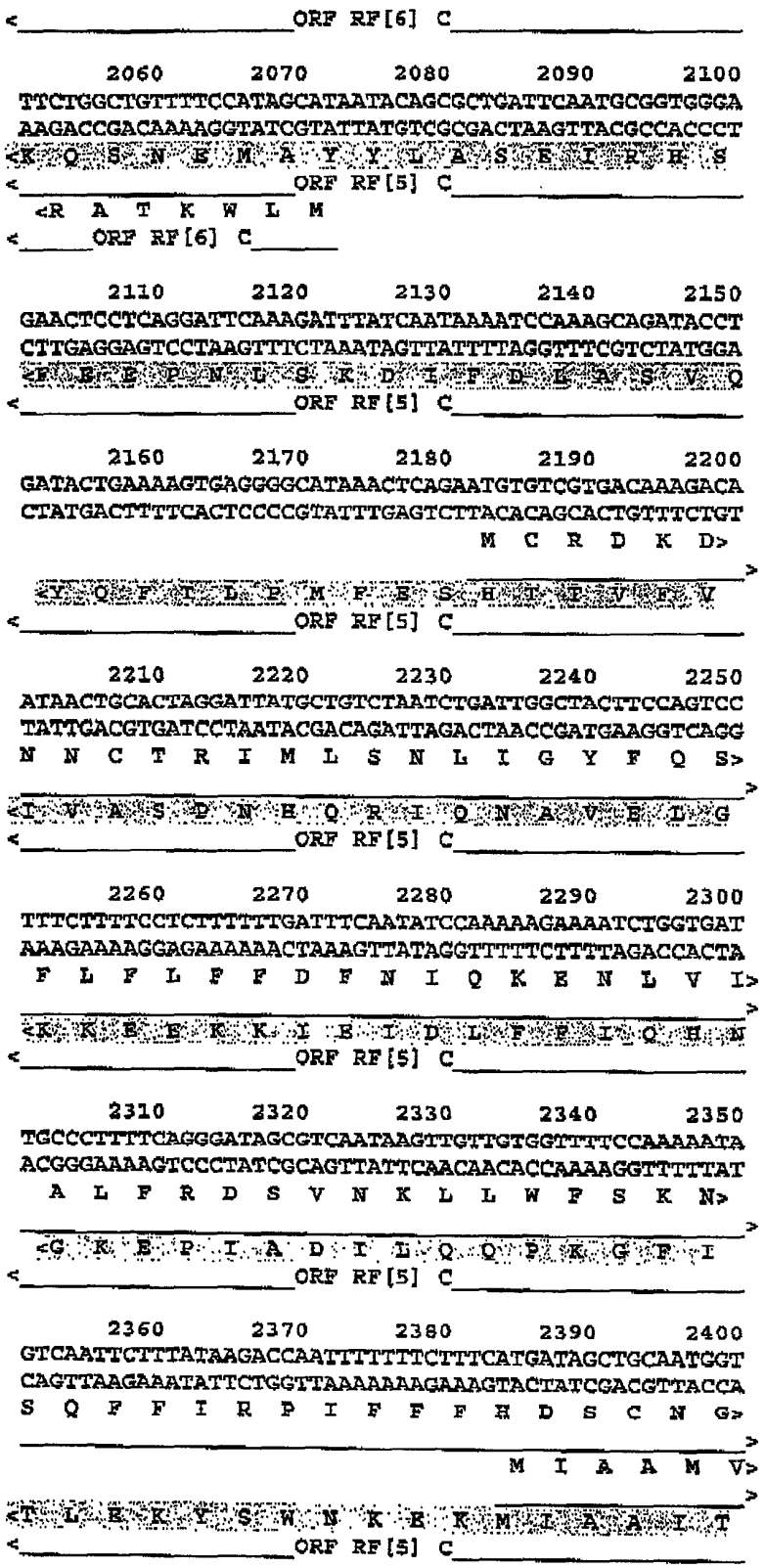





N Q I N A L H G I N Y R T D Q L Q
ORF RF[4]
860 870 880 890 900
ATGTCCTGTTTTTTCTAAGATTTGATGGTAAATCTTTTCAATACTAGCTA
TACAGGACAAAAAAGATTCTAAACTACCATTTAGAAAAGTTATGATCGAT
H G T K E L I Q H Y I K E I S A L
ORF RF[4]
910 920 930 940 950
AATCTTTATTTTCAATGCCTAATCTGAGGCTAGTTAAAATATTCAGATAA
TTAGAAATAAAAGTTACGGATTAGACTCCGATCAATTTTATAAGTCTATT
D K N E I G L R L S I I I N I Y
ORF RF[4]
960 970 980 990 1000
TCATGGCGGAAACTTCGAATATCCTTGTAAAGAGATTCTATTTGCTGACT
AGTACCGCCTTTGAAGCTTATAGGAACATTTCTCTAAGATAAACGACTGA
M A E T S N I L V K R F Y L L T>
D H R E S R I D K Y L S E I Q Q S
ORF RF[4]
1010 1020 1030 1040 1050
ATACTGGGTGATATTTCGAATCTGAGCTTCCTTTGTGCCATTATCTCAT
TATGACCCACTATAAAGCTTAGACTCGAAGGAAAACACGGTAATAGAGTA
I L G D I S N L S F L L C H Y L I>
V Q T I N R I Q A E K Q A M I E N
ORF RF[4]
1060 1070 1080 1090 1100
TTTGAACCTTTTGTTTGGTATATTGGCTTAAAAATGAGATCAGAATCAAA
AAACTTGGAAAACAAACCATATAACCGAATTTTTACTCTAGTCTTAGTTT
L N L L F G I L A>
Q V K Q K T Y Q S L F S I L I L
ORF RF[4]
1110 1120 1130 1140 1150
AATAAAATAAGATACACAATAACGACAAATTTACGAAATTTTAAAGTCGG
TTATTTTATTCTATCTGTTATTGCTGTTTAAATGCTTTAAAATTTCAGCC
F L I L Y V D V V E K R F K L T P
ORF RF[4]
1160 1170 1180 1190 1200
TATCACATTATAACTCTCTATAACATACAATACCTGTATTAAAAGGTAGT
ATAGTGTAATATTGAGAGATATTGTATGTTATGGACATAATTTTCCATCA
I V N Y S E I V Y L V Q I L L Y Y
ORF RF[4]
1210 1220 1230 1240 1250
ATAGAAGCATAGTAATATTCATTGGAATCAAGCGTTTTTTGACCTTCATC
TATCTTCGTATCATTATAAGTAACCTTAGTTCGCAAAAAACTGGAAGTAG
L L M T I N M E I L R K K V K M
ORF RF[4]

```
        1260      1270      1280      1290      1300
 TTGGTCAAACTATCTTTAAGTCGACCAATATCAACATTGAACACACTGAG
 AACCAGTTTGATAGAAATTCAGCTGGTTATAGTTGTAACTTGTGTGACTC
<K  T  L  S  D  K  L  R  G  L  D  V  N  F  V  S  L
<____________________ORF RF[4] C____________________

        1310      1320      1330      1340      1350
 AAAGAGGTAACTTAGCGCTATTCCTGCAAACTCGATCATATAGGTGGTTA
 TTTCTCCATTGAATCGCGATAAGGACGTTTGAGCTAGTATATCCACCAAT
<F  L  Y  S  L  A  I  G  A  F  E  I  M  Y  T  T  I
<____________________ORF RF[4] C____________________

        1360      1370      1380      1390      1400
 TAATGCTACTGCCCATTACAATTCCTTGAGTTCCATCCAAGATAAAGAAT
 ATTACGATGACGGGTAATGTTAAGGAACTCAAGGTAGGTTCTATTTCTTA
  <I  S  S  G  M  V  I  G  Q  T  G  D  L  I  F  F
<____________________ORF RF[4] C____________________
                    <L  E  K  L  E  M  W  S  L  S  Y
                   <________________________________

        1410      1420      1430      1440      1450
 ATGATTGCCCGCCTAAACAAGTCAGAACTGGCAACAGGCAGCAGACCATA
 TACTAACGGGCGGATTTGTTCAGTCTTGACCGTTGTCCGTCGTCTGGTAT
<I  I  A  R  R  F  L  D  S  S  A  V  P  L  L  G  Y
<____________________ORF RF[4] C____________________
  <S  Q  G  G  L  C  T  L  V  P  L  L  C  C  V  M
<___________________________________________________

        1460      1470      1480      1490      1500
 AAATATATTTAGAGAAAGACTATTCTGTCTATTAAGATAAATTGATAAAG
 TTTATATAAATCTCTTTCTGATAAGACAGATAATTCTATTTAACTATTTC
<F  I  N  L  S  L  S  N  Q  R  N  L  Y  I  S  L  A
<____________________ORF RF[4] C____________________

        1510      1520      1530      1540      1550
 CTATAAAATAAAGAGGCTCTGCAGGATAAAACAGGTTTACGTTCACCATC
 GATATTTTATTTCTCCGAGACGTCCTATTTTGTCCAAATGCAAGTGGTAG
  <I  F  Y  L  P  E  A  P  Y  F  L  N  V  N  V  M
<____________________ORF RF[4] C____________________

        1560      1570      1580      1590      1600
 GTAACAGCAATCATTATCAGAAAATTGCTTATCGAAAAAAGAGTTAATTC
 CATTGTCGTTAGTAATAGTCTTTTAACGAATAGCTTTTTTCTCAATTAAG
<T  V  A  I  M  I  L  F  N  S  I  S  F  L  T  L  E
<____________________ORF RF[4] C____________________

        1610      1620      1630      1640      1650
 CTTTTTCGATAAAGTGACATTACTTACCTTAGAAAATAGAAACAAGAGAA
 GAAAAAGCTATTTCACTGTAATGAATGGAATCTTTTATCTTTGTTCTCTT
<K  K  S  L  T  V  N  S  V  K  S  F  L  F  L  L  F
<____________________ORF RF[4] C____________________

        1660      1670      1680      1690      1700
 ATAGAACGGTTAGATAAGTTAATAAACCATTTGAAAGTATCATTAAGGCT
 TATCTTGCCAATCTATTCAATTATTTGGTAAACTTTCATAGTAATTCCGA
  <L  V  T  L  Y  T  L  L  G  N  S  L  I  M  L  A
<____________________ORF RF[4] C____________________
```

```
        1710      1720      1730      1740      1750
TCATTCATTTTGCTCTCCTTTGATCAGCAATCACAGCTCTCAGTTTTGTT
AGTAAGTAAAACGAGAGGAAACTAGTCGTTAGTGTCGAGAGTCAAAACAA
<E  N  M
<________
    <K  A  R  R  Q  D  A  I  V  A  R  L  K  T
    <________________ORF RF[5] C______________

        1760      1770      1780      1790      1800
AACTTAGTTCGTGAAATAAGACAAGACTTATTATTTCGAAAATAGGCCAA
TTGAATCAAGCACTTTATTCTGTTCTGAATAATAAAGCTTTTATCCGGTT
<L  K  T  R  S  I  L  C  S  K  N  N  R  F  Y  A  L
<____________________ORF RF[5] C__________________

        1810      1820      1830      1840      1850
GCGTTTTTTCCGATCAATACGGGTAATATTGGCAGGATTGACAATAAAAG
CGCAAAAAAGGCTAGTTATGCCCATTATAACCGTCCTAACTGTTATTTTC
<R  K  K  R  D  I  R  T  I  N  A  P  N  V  I  F  S
<____________________ORF RF[5] C__________________

        1860      1870      1880      1890      1900
AGCGATGGCACTGAAAAAGTCTCTTATCCATTTTAACAATGTCAGTCATA
TCGCTACCGTGACTTTTTCAGAGAATAGGTAAAATTGTTACAGTCAGTAT
     M  A  L  K  K  S  L  I  H  F  N  N  V  S  H>
  ______________________ORF RF[1] _______________>
  <R  H  C  Q  F  L  R  K  D  M  K  V  I  D  T  M
<____________________ORF RF[5] C__________________
                                                <V
                                                <__

        1910      1920      1930      1940      1950
CTGCCGTAGAATTCAATCCGTTCATCATAAGTATAAAGGCAGAGCTTATG
GACGGCATCTTAAGTTAGGCAAGTAGTATTCATATTTCCGTCTCGAATAC
T  A  V  E  F  N  P  F  I  I  S  I  K  A  E  L  M>
_____________________ORF RF[1] __________________>
<S  G  Y  F  E  I  R  E  D  Y  T  Y  L  C  L  K  H
<____________________ORF RF[5] C__________________
 <A  T  S  N  L  G  N  M  M  L  I  F  A  S  S  I
<____________________ORF RF[6] C__________________

        1960      1970      1980      1990      2000
GGCTGTTGAAGATGTTTCAAAATACAGAATCTCAGCAAAAGGGACCTGAA
CCGACAACTTCTACAAAGTTTTATGTCTTAGAGTCGTTTTCCCTGGACTT
 G  C>
_____>
<A  T  S  S  T  E  F  Y  L  I  E  A  F  P  V  Q  F
<____________________ORF RF[5] C__________________
<P  Q  Q  L  H  K  L  I  C  F  R  L  L  L  S  R  F
<____________________ORF RF[6] C__________________

        2010      2020      2030      2040      2050
ACTGAGTTTCAGATGAATGGAAAATAAAAAGTTCCTCTGATTGACCATTC
TGACTCAAAGTCTACTTACCTTTTATTTTTCAAGGAGACTAACTGGTAAG
 <Q  T  E  S  S  H  F  I  F  L  E  E  S  Q  G  N
<____________________ORF RF[5] C__________________
<S  L  K  L  H  I  S  F  L  F  N  R  Q  N  V  M  R
```


ORF RF[6] C
2060 2070 2080 2090 2100
TTCTGGCTGTTTTCCATAGCATAATACAGCGCTGATTCAATGCGGTGGGA
AAGACCGACAAAAGGTATCGTATTATGTCGCGACTAAGTTACGCCACCCT
K Q S N E M A Y Y L A S E I R E S
ORF RF[5] C
R A T K W L M
ORF RF[6] C
2110 2120 2130 2140 2150
GAACTCCTCAGGATTCAAAGATTTATCAATAAAATCCAAAGCAGATACCT
CTTGAGGAGTCCTAAGTTTCTAAATAGTTATTTTAGGTTTCGTCTATGGA
E L E P N L S K D I T D E A S V Q
ORF RF[5] C
2160 2170 2180 2190 2200
GATACTGAAAAGTGAGGGGCATAAACTCAGAATGTGTCGTGACAAAGACA
CTATGACTTTTCACTCCCCGTATTTGAGTCTTACACAGCACTGTTTCTGT
M C R D K D>
Y Q F T L P M F E S H I E V F V
ORF RF[5] C
2210 2220 2230 2240 2250
ATAACTGCACTAGGATTATGCTGTCTAATCTGATTGGCTACTTCCAGTCC
TATTGACGTGATCCTAATACGACAGATTAGACTAACCGATGAAGGTCAGG
N N C T R I M L S N L I G Y F Q S>
I V A S D N H Q R I Q N A V E L G
ORF RF[5] C
2260 2270 2280 2290 2300
TTTCTTTTCCTCTTTTTTGATTTCAATATCCAAAAAGAAAATCTGGTGAT
AAAGAAAAGGAGAAAAAACTAAAGTTATAGGTTTTTCTTTTAGACCACTA
F L F L F F D F N I Q K E N L V I>
K K E E K K I E I D L F E I Q H N
ORF RF[5] C
2310 2320 2330 2340 2350
TGCCCTTTTCAGGGATAGCGTCAATAAGTTGTTGTGGTTTTCCAAAAATA
ACGGGAAAAGTCCCTATCGCAGTTATTCAACAACACCAAAAGGTTTTTAT
A L F R D S V N K L L W F S K N>
G K E P I A D T L Q Q P K G F I
ORF RF[5] C
2360 2370 2380 2390 2400
GTCAATTCTTTATAAGACCAATTTTTTTCTTTCATGATAGCTGCAATGGT
CAGTTAAGAAATATTCTGGTTAAAAAAAGAAAGTACTATCGACGTTACCA
S Q F F I R P I F F F H D S C N G>
M I A A M V>
T L E K Y S W N K E K M I A A I T
ORF RF[5] C

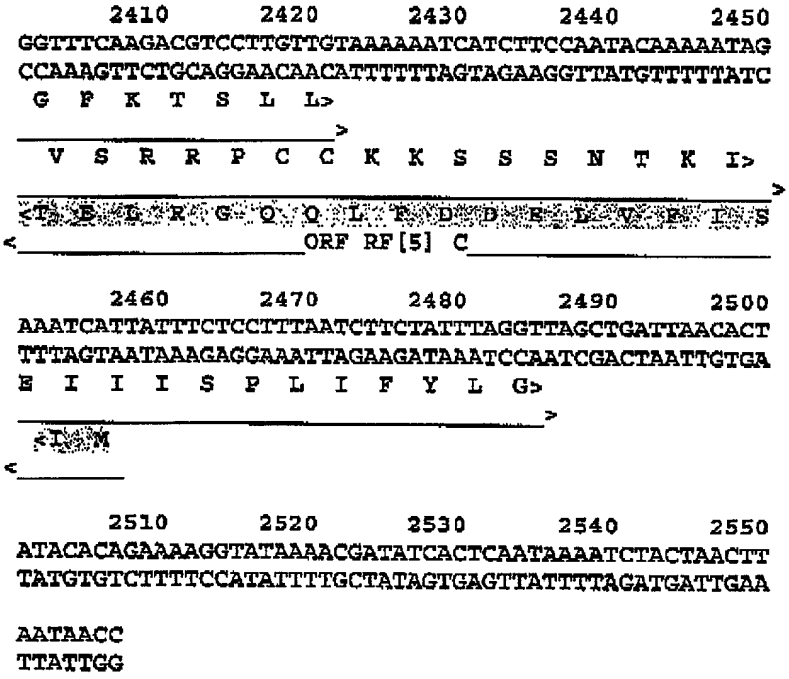


Figure 10
The comX nucleotide sequence, amino acid sequence, and its local region with 100bp included both upstream and downstream (promoter is upstream).

> S. mutans comX gene

ATGGAAGAAGATTTTGAAATTGTTTTTAATAAGGTTAAGCCAATTGTATGGAAATTAAG
CCGTTATTACTTTATTAAAATGTGGACTCGTGAAGATTGGCAACAAGAGGGAATGTTGA
TTTTGCACCAATTATTAAGGGAACATCCAGAATTAGAAGAGGATGATACAAAATTGTAT
ATCTATTTTAAGACACGTTTTTCTAATTACATTAAAGATGTTTTGCGTCAGCAAGAAAG
TCAGAAACGTCGTTTTAATAGAATGTCTTATGAAGAAGTCGGTGAGATTGAACACTGTT
TGTCAAGTGGCGGTATGCAATTGGATGAATATATTTTATTTCGTGATAGTTTGCTTGCA
TATAAACAAGGTCTGAGTACTGAAAAGCAAGAGCTGTTTGAGCGCTTGGTAGCAGGAGA
GCACTTTTTGGGAAGGCAAAGTATGCTGAAAGATTTACGTAAAAAATTAAGTGATTTTA
AGGAAAAA

> S. mutans ComX protein

MEEDFEIVFNKVKPIVWKLSRYYFIKMWTREDWQQEGMLILHQLLREHPELEEDDTKLY
IYFKTRFSNYIKDVLRQQESQKRRFNRMSYEEVGEIEHCLSSGGMQLDEYILFRDSLLA
YKQGLSTEKQELFERLVAGEHFLGRQSMLKDLRKKLSDFKEK

> S. mutans comX gene local region

GTAAATAAAACAGCCAGTTAAGATGGGACATTTATGTCCTGTTCTTAAAGTCTTTTTCG
TTTTATAATAATTTTATTATAAAAGGAGGTCATCGTAATAGATGGAAGAAGATTTTGAA
ATTGTTTTTAATAAGGTTAAGCCAATTGTATGGAAATTAAGCCGTTATTACTTTATTAA
AATGTGGACTCGTGAAGATTGGCAACAAGAGGGAATGTTGATTTTGCACCAATTATTAA
GGGAACATCCAGAATTAGAAGAGGATGATACAAAATTGTATATCTATTTTAAGACACGT
TTTTCTAATTACATTAAAGATGTTTTGCGTCAGCAAGAAAGTCAGAAACGTCGTTTTAA
TAGAATGTCTTATGAAGAAGTCGGTGAGATTGAACACTGTTTGTCAAGTGGCGGTATGC
AATTGGATGAATATATTTTATTTCGTGATAGTTTGCTTGCATATAAACAAGGTCTGAGT
ACTGAAAAGCAAGAGCTGTTTGAGCGCTTGGTAGCAGGAGAGCACTTTTTGGGAAGGCA
AAGTATGCTGAAAGATTTACGTAAAAAATTAAGTGATTTTAAGGAAAAATAGTTAAAAA
GGGAAAGAATGGAACATGTGATTGTACCATTCTTTTTGGTTGAAAATTAAGAAAAGTTA
TTATAAATTATTGGTTTAACATGCCATATTA

Figure 11.
The comA and comB nucleotide and amino acid sequences. ComA and ComB are the components of the CSP exporter.

➢ S. mutans comA gene

ATGAAACAAGTTATTTATGTTGTTTTAATCGTCATAGCCGTTAACATTCTCTTAGAGAT
TATCAAAAGAGTAACAAAAAGGGGAGGGACAGTTTCGTCATCTAATCCTTTACCAGATG
GGCAGTCTAAGTTGTTTTGGCGCAGACATTATAAGCTAGTACCTCAGATTGATACCAGA
GACTGTGGGCCGGCAGTGCTGGCATCTGTTGCAAAGCATTACGGATCTAATTACTCTAT
CGCTTATCTGCGGGAACTCTCAAAGACTAACAAGCAGGGAACAACAGCTCTTGGCATTG
TTGAAGCTGCTAAAAAGTTAGGCTTTGAAACACGCTCTATCAAGGCGGATATGACGCTT
TTTGATTATAATGATTTGACCTATCCTTTTATCGTCCATGTGATTAAAGGAAAACGTCT
GCAGCATTATTATGTCGTCTATGGCAGCCAGAATAATCAGCTGATTATTGGAGATCCTG
ATCCTTCAGTTAAGGTGACTAGGATGAGTAAGGAACGCTTTCAATCAGAGTGGACAGGC
CTTGCAATTTTCCTAGCTCCTCAGCCTAACTATAAGCCTCATAAAGGTGAAAAAAATGG
TTTGTCTAATTTCTTCCCGTTGATCTTTAAGCAGAAAGCTTTGATGACTTATATTATCA
TAGCTAGCTTGATTGTGACGCTCATTGATATTGTCGGATCATACTATCTCCAAGGAATA
TTGGACGAGTACATTCCTGATCAGCTGATTTCAACTTTAGGAATGATTACGATTGGTCT
GATAATAACCTATATTATCCAGCAGGTCATGGCTTTTGCAAAAGAATACCTCTTGGCCG
TACTCAGTTTGCGTTTAGTCATTGATGTTATCCTGTCTTATATCAAACATATTTTTACG
CTTCCTATGTCTTTCTTTGCGACAAGGCGAACAGGAGAAATCACGTCTCGTTTTACAGA
TGCCAATCAGATTATTGATGCTGTAGCGTCAACCATCTTTTCAATCTTTTTAGATATGA
CTATGGTAATTTTGGTTGGTGGGGTTTTGTTGGCGCAAAACAATAACCTTTTCTTTCTA
ACCTTGCTCTCCATTCCGATTTATGCCATCATTATTTTGCTTTCTTGAAACCCTTTGA
GAAAATGAATCACGAAGTGATGGAAAGCAATGCTGTGGTAAGTTCTTCTATCATTGAAG
ATATCAATGGGATGGAAACCATTAAATCACTCACAAGTGAGTCCGCTCGTTATCAAAAC
ATTGATAGTGAATTTGTTGATTATTTGGAGAAAAACTTTAAGCTACACAAGTATAGTGC
CATTCAAACCGCATTAAAAAGCGGTGCTAAGCTTATCCTCAATGTTGTCATTCTCTGGT
ATGGCTCTCGTCTAGTTATGGATAATAAAATCTCAGTTGGTCAGCTTATCACCTTTAAT
GCTTTGCTGTCTTATTTCTCAAATCCAATTGAAAATATTATCAATCTGCAATCCAAACT
GCAGTCAGCTCGCGTTGCCAATACACGTCTTAATGAGGTCTATCTTGTCGAATCTGAAT
TTGAAAAAGACGGCGATTTATCAGAAAATAGCTTTTTAGATGGTGATATTTCGTTTGAA
AATCTTTCTTATAAATATGGATTTGGGCGAGATACCTTATCAGATATTAATTTATCAAT
CAAAAAAGGCTCCAAGGTCAGTCTAGTTGGAGCCAGTGGTTCTGGTAAAACAACTTTGG
CTAAACTGATTGTCAATTTCTACGAGCCTAACAAGGGGATTGTTCGAATCAATGGCAAT
GATTTAAAAGTTATTGATAAGACAGCTTTGCGGCGGCATATTAGCTATTTGCCGCAACA
GGCCTATGTTTTTAGTGGCTCTATTATGGATAATCTCGTTTTAGGAGCTAAAGAAGGAA
CGAGTCAGGAAGACATTATTCGTGCTTGTGAAATTGCTGAAATCCGCTCGGACATTGAA
CAAATGCCTCAGGGCTATCAGACAGAGTTATCAGATGGTGCCGGTATTTCTGGCGGTCA
AAAACAGCGGATTGCTTTAGCTAGGGCCTTATTAACACAGGCACCGGTTTTGATTCTGG
ATGAAGCCACCAGCAGTCTTGATATTTTGACAGAAAAGAAAATTATCAGCAATCTCTTA
CAGATGACGGAGAAAACAATAATTTTTGTTGCCCACCGCTTAAGCATTTCACAGCGTAC
TGACGAAGTCATTGTCATGGATCAGGGAAAAATTGTTGAACAAGGCACTCATAAGGAAC
TTTTAGCTAAGCAAGGTTTCTATTATAACCTGTTTAAT

➢ S. mutans ComA protein

MKQVIYVVLIVIAVNILLEIIKRVTKRGGTVSSSNPLPDGQSKLFWRRHYKLVPQIDTR
DCGPAVLASVAKHYGSNYSIAYLRELSKTNKQGTTALGIVEAAKKLGFETRSIKADMTL
FDYNDLTYPFIVHVIKGKRLQHYYVVYGSQNNQLIIGDPDPSVKVTRMSKERFQSEWTG
LAIFLAPQPNYKPHKGEKNGLSNFFPLIFKQKALMTYIIIASLIVTLIDIVGSYYLQGI
LDEYIPDQLISTLGMITIGLIITYIIQQVMAFAKEYLLAVLSLRLVIDVILSYIKHIFT
LPMSFFATRRTGEITSRFTDANQIIDAVASTIFSIFLDMTMVILVGGVLLAQNNNLFFL
TLLSIPIYAIIIFAFLKPFEKMNHEVMESNAVVSSSIIEDINGMETIKSLTSESARYQN
IDSEFVDYLEKNFKLHKYSAIQTALKSGAKLILNVVILWYGSRLVMDNKISVGQLITFN
ALLSYFSNPIENIINLQSKLQSARVANTRLNEVYLVESEFEKDGDLSENSFLDGDISFE
NLSYKYGFGRDTLSDINLSIKKGSKVSLVGASGSGKTTLAKLIVNFYEPNKGIVRINGN
DLKVIDKTALRRHISYLPQQAYVFSGSIMDNLVLGAKEGTSQEDIIRACEIAEIRSDIE
QMPQGYQTELSDGAGISGGQKQRIALARALLTQAPVLILDEATSSLDILTEKKIISNLL
QMTEKTIIFVAHRLSISQRTDEVIVMDQGKIVEQGTHKELLAKQGFYYNLFN

➢ S. mutans comB gene

ATGGATCCTAAATTTTTACAAAGTGCAGAATTTTATAGGAGACGCTATCATAATTTTGC
GACACTATTAATTGTTCCTTTGGTCTGCTTGATTATCTTCTTGGTCATATTCCTTTGTT
TTGCTAAAAAAGAAATTACAGTGATTTCTACTGGTGAAGTTGCACCAACAAAGGTTGTA
GATGTTATCCAATCTTACAGTGACAGTTCAATCATTAAAAATAATTTAGATAATAATGC
AGCTGTTGAGAAGGGAGACGTTTTAATTGAATATTCAGAAAATGCCAGTCCAAACCGTC
AGACTGAACAAAAGAATATTATAAAAGAAAGACAAAAACGAGAAGAGAAGGAAAAGAAA
AAACACCAAAAGAGCAAGAAAAAGAAGAAGTCTAAGAGCAAGAAAGCTTCCAAAGATAA
GAAAAAGAAATCGAAAGACAAGGAAAGCAGCTCTGACGATGAAAATGAGACAAAAAAGG
TTTCGATTTTTGCTTCAGAAGATGGTATTATTCATACCAATCCCAAATATGATGGTGCC
AATATTATTCCGAAGCAAACCGAGATTGCTCAAATCTATCCTGATATTCAAAAAACAAG
AAAAGTGTTAATCACCTATTATGCTTCTTCTGATGATGTTGTTTCTATGAAAAAGGGGC
AAACCGCTCGTCTTTCCTTGGAAAAAAAGGGAAATGACAAGGTTGTTATTGAAGGAAAA
ATTAACAATGTCGCTTCATCAGCAACTACTACTAAAAAAGGAAATCTCTTTAAGGTTAC
TGCCAAAGTAAAGGTTTCTAAGAAAAATAGCAAACTCATCAAGTATGGTATGACAGGCA
AGACAGTCACTGTCATTGATAAAAAGACTTATTTTGATTATTTCAAAGATAAATTACTG
CATAAAATGGATAAT

➢ S. mutans ComB protein

MDPKFLQSAEFYRRRYHNFATLLIVPLVCLIIFLVIFLCFAKKEITVISTGEVAPTKVV
DVIQSYSDSSIIKNNLDNNAAVEKGDVLIEYSENASPNRQTEQKNIIKERQKREEKEKK
KHQKSKKKKKSKSKKASKDKKKKSKDKESSSDDENETKKVSIFASEDGIIHTNPKYDGA
NIIPKQTEIAQIYPDIQKTRKVLITYYASSDDVVSMKKGQTARLSLEKKGNDKVVIEGK
INNVASSATTTKKGNLFKVTAKVKVSKKNSKLIKYGMTGKTVTVIDKKTYFDYFKDKLL
HKMDN

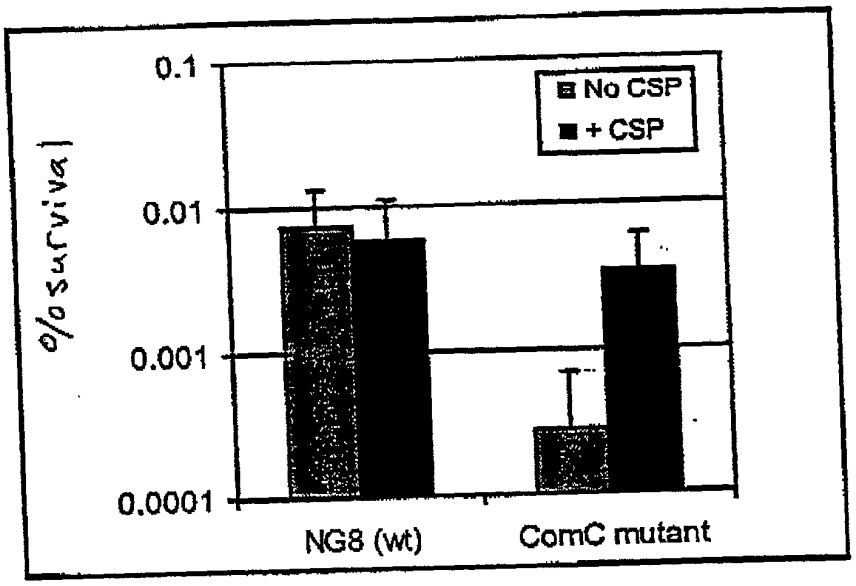


Figure 12